



OIPE

3 ENTERED

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/10/076,157

DATE: 04/16/2002
 TIME: 14:33:26

Input Set : A:\EP.txt
 Output Set: N:\CRF3\04162002\J076157.raw

3 <110> APPLICANT: Pompejus, Markus
 4 Suelberger, Harald
 5 Joeffken, Hans Wolfgang
 6 Doval, Jose Luis Revuelta
 7 Jimenez, Alberto;
 8 Garcia, Maria Angeles Santos
 10 <120> TITLE OF INVENTION: Genes of purine biosynthesis from Ashbya Gossypii and the
 use thereof
 11 in
 12 microbial riboflavin synthesis
 14 <130> FILE REFERENCE: 48684DIV
 16 <140> CURRENT APPLICATION NUMBER: US 10/076,157
 17 <141> CURRENT FILING DATE: 2002-02-15
 19 <150> PRIOR APPLICATION NUMBER: US 09/212,247
 20 <151> PRIOR FILING DATE: 1998-12-16
 22 <160> NUMBER OF SEQ ID NOS: 21
 24 <170> SOFTWARE: WordPerfect v. 6.1
 26 <210> SEQ ID NO: 1
 27 <211> LENGTH: 1911
 28 <212> TYPE: DNA
 29 <213> ORGANISM: Ashbya gosypii
 31 <220> FEATURE:
 32 <221> NAME/KEY: CDS
 33 <222> LOCATION: 626..1582
 35 <400> SEQUENCE: 1
 37 ggtagtcgct catcgacaga cacaatcgcg tgttctctct gaatcgtcca ttgggtgtca 60
 39 gcatcctgat cgcgggcgga tggaatgggt aatcattagg aaacaccaat gtcccatggg 120
 41 attgtccgctc ctctgatggg gtctcaggag gaccctgat cacgtagtgc cacaccagga 180
 43 tattgtcttc ctttgggtgt gccacgatgt agggcggggg gttctcgggc atcattttgt 240
 45 actcctttga gagccgcttg tacgcctgtc ttgatgccat cttgcctact attagtttct 300
 47 caccacttcc cgccaaacaa tctgcacttt acgagcgcta tctatccctc gggtcgctct 360
 49 agttgattat tggcgaaact gatagttcag gtacttccat gatgcgggtca tatccacgta 420
 51 tgtgatcacg tgatcatcag ccatgctgcc agctcacggg cctgcctaca ctattggagg 480
 53 ctctgtgagt catgatttat tgcatatcaa gccagatag tcgttgggga tactaccgtt 540
 55 gcgcgatga gctccgatat taagttgtag ccaaaaattt taacggatga cttcttaaca 600
 57 gttattgacg ccgcaatcct acgcc atg tcc aat agc ata aag ctg cta 652
 58 Met Ser Ser Asn Ser Ile Lys Leu Leu
 59 1 5
 61 gca ggt aac tcg cac ccg gac cta gct gag aag gtc tcc gtt cgc cta 700
 62 Ala Gly Asn Ser His Pro Asp Leu Ala Glu Lys Val Ser Val Arg Leu
 63 10 15 20 25
 65 ggt gta cca ctt tcg aag att gga gtg tat cac tac tct aac aaa gag 748
 66 Gly Val Pro Leu Ser Lys Ile Gly Val Tyr His Tyr Ser Asn Lys Glu
 67 30 35 40

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/076,157

DATE: 04/16/2002

TIME: 14:33:26

Input Set : A:\EP.txt

Output Set: N:\CRF3\04162002\J076157.raw

69	acg tca gtt act atc ggc gaa agt atc cgt gat gaa gat gtc tac atc	796
70	Thr Ser Val Thr Ile Gly Glu Ser Ile Arg Asp Glu Asp Val Tyr Ile	
71	45 50 55	
73	atc cag aca gga acg ggg gag cag gaa atc aac gac ttc ctc atg gaa	844
74	Ile Gln Thr Gly Thr Gly Glu Gln Glu Ile Asn Asp Phe Leu Met Glu	
75	60 65 70	
77	ctg ctc atc atg atc cat gcc tgc cgg tca gcc tct gcg cgg aag atc	892
78	Leu Leu Ile Met Ile His Ala Cys Arg Ser Ala Ser Ala Arg Lys Ile	
79	75 80 85	
81	aca gcg gtt ata cca aac ttc cct tac gca aga caa gac aaa aag gac	940
82	Thr Ala Val Ile Pro Asn Phe Pro Tyr Ala Arg Gln Asp Lys Lys Asp	
83	90 95 100 105	
85	aag tcg cga gca ccg ata act gcc aag ctg gtg gcc aag atg cta gag	988
86	Lys Ser Arg Ala Pro Ile Thr Ala Lys Leu Val Ala Lys Met Leu Glu	
87	110 115 120	
89	acc gcg ggg tgc aac cac gtt atc acg atg gat ttg cac gcg tct caa	1036
90	Thr Ala Gly Cys Asn His Val Ile Thr Met Asp Leu His Ala Ser Gln	
91	125 130 135	
93	att cag ggt ttc ttc cac att cca gtg gac aac cta tat gca gag ccg	1084
94	Ile Gln Gly Phe Phe His Ile Pro Val Asp Asn Leu Tyr Ala Glu Pro	
95	140 145 150	
97	aac atc ctg cac tac atc caa cat aat gtg gac ttc cag aat agt atg	1132
98	Asn Ile Leu His Tyr Ile Gln His Asn Val Asp Phe Gln Asn Ser Met	
99	155 160 165	
101	ttg gtc gcg cca gac gcg ggg tcg gcg aag cgc acg tcg acg ctt tcg	1180
102	Leu Val Ala Pro Asp Ala Gly Ser Ala Lys Arg Thr Ser Thr Leu Ser	
103	170 175 180 185	
105	gac aag ctg aat ctc aac ttc gcg ttg atc cac aaa gaa cgg cag aag	1228
106	Asp Lys Leu Asn Leu Asn Phe Ala Leu Ile His Lys Glu Arg Gln Lys	
107	190 195 200	
109	gcg aac gag gtc tcg cgg atg gtg ttg gtg ggt gat gtc gcc gac aag	1276
110	Ala Asn Glu Val Ser Arg Met Val Leu Val Gly Asp Val Ala Asp Lys	
111	205 210 215	
113	tcc tgt att att gta gac gac atg gcg gac acg tgc gga acg cta gtg	1324
114	Ser Cys Ile Ile Val Asp Asp Met Ala Asp Thr Cys Gly Thr Leu Val	
115	220 225 230	
117	aag gcc act gac acg ctg atc gaa aat tgt gcg aaa gaa gtg att gcc	1372
118	Lys Ala Thr Asp Thr Leu Ile Glu Asn Cys Ala Lys Glu Val Ile Ala	
119	235 240 245	
121	att gtg aca cac ggt ata ttt tct ggc ggc gcc cgc gag aag ttg cgc	1420
122	Ile Val Thr His Gly Ile Phe Ser Gly Gly Ala Arg Glu Lys Leu Arg	
123	250 255 260 265	
125	aac agc aag ctg gca cgg atc gta agc aca aat acg gtg cca gtg gac	1468
126	Asn Ser Lys Leu Ala Arg Ile Val Ser Thr Asn Thr Val Pro Val Asp	
127	270 275 280	
129	ctc aat cta gat atc tac cac caa att gac att agt gcc att ttg gcc	1516
130	Leu Asn Leu Asp Ile Tyr His Gln Ile Asp Ile Ser Ala Ile Leu Ala	
131	285 290 295	
133	gag gca att aga agg ctt cac aac ggg gaa agt gtg tcg tac ctg ttc	1564

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/076,157

DATE: 04/16/2002

TIME: 14:33:26

Input Set : A:\EP.txt

Output Set: N:\CRF3\04162002\J076157.raw

```

134 Glu Ala Ile Arg Arg Leu His Asn Gly Glu Ser Val Ser Tyr Leu Phe
135          300          305          310
137 aat aac gct gtc atg tagtgctgctc agtggcagat gcatgatcgc tggcctaatt      1619
138 Asn Asn Ala Val Met
139          315
141 atctgtgtaa gttgatacaa tgcagtaaat acagtacata aaactgaatg tttttcactt      1679
143 aggggtgctt tggtgttctg atagcgtgtg tgcgaatttg gaggtgaaag ttgaacatca      1739
145 cgtaatgaat acaaacaaga ttgcacatta ggaaaagcga taaattattt attatttgca      1799
147 actggccttt gagcgtttta gcctgaacat ttttgcctt ttgtttgacc gtaccgttat      1859
149 cactcgtcct tatatatggc tatecttctc ttccggaact tcttcgagcg ta      1911
154 <210> SEQ ID NO: 2
155 <211> LENGTH: 318
156 <212> TYPE: PRT
157 <213> ORGANISM: Ashbya gosypii
159 <400> SEQUENCE: 2
161 Met Ser Ser Asn Ser Ile Lys Leu Leu Ala Gly Asn Ser His Pro Asp
162 1          5          10          15
164 Leu Ala Glu Lys Val Ser Val Arg Leu Gly Val Pro Leu Ser Lys Ile
165          20          25          30
167 Gly Val Tyr His Tyr Ser Asn Lys Glu Thr Ser Val Thr Ile Gly Glu
168          35          40          45
170 Ser Ile Arg Asp Glu Asp Val Tyr Ile Ile Gln Thr Gly Thr Gly Glu
171          50          55          60
173 Gln Glu Ile Asn Asp Phe Leu Met Glu Leu Leu Ile Met Ile His Ala
174 65          70          75          80
176 Cys Arg Ser Ala Ser Ala Arg Lys Ile Thr Ala Val Ile Pro Asn Phe
177          85          90          95
179 Pro Tyr Ala Arg Gln Asp Lys Lys Asp Lys Ser Arg Ala Pro Ile Thr
180          100          105          110
182 Ala Lys Leu Val Ala Lys Met Leu Glu Thr Ala Gly Cys Asn His Val
183          115          120          125
185 Ile Thr Met Asp Leu His Ala Ser Gln Ile Gln Gly Phe Phe His Ile
186          130          135          140
188 Pro Val Asp Asn Leu Tyr Ala Glu Pro Asn Ile Leu His Tyr Ile Gln
189 145          150          155          160
191 His Asn Val Asp Phe Gln Asn Ser Met Leu Val Ala Pro Asp Ala Gly
192          165          170          175
194 Ser Ala Lys Arg Thr Ser Thr Leu Ser Asp Lys Leu Asn Leu Asn Phe
195          180          185          190
197 Ala Leu Ile His Lys Glu Arg Gln Lys Ala Asn Glu Val Ser Arg Met
198          195          200          205
200 Val Leu Val Gly Asp Val Ala Asp Lys Ser Cys Ile Ile Val Asp Asp
201          210          215          220
203 Met Ala Asp Thr Cys Gly Thr Leu Val Lys Ala Thr Asp Thr Leu Ile
204 225          230          235          240
206 Glu Asn Cys Ala Lys Glu Val Ile Ala Ile Val Thr His Gly Ile Phe
207          245          250          255
209 Ser Gly Gly Ala Arg Glu Lys Leu Arg Asn Ser Lys Leu Ala Arg Ile
210          260          265          270

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/076,157

DATE: 04/16/2002

TIME: 14:33:26

Input Set : A:\EP.txt

Output Set: N:\CRF3\04162002\J076157.raw

```

212 Val Ser Thr Asn Thr Val Pro Val Asp Leu Asn Leu Asp Ile Tyr His
213          275          280          285
215 Gln Ile Asp Ile Ser Ala Ile Leu Ala Glu Ala Ile Arg Arg Leu His
216          290          295          300
218 Asn Gly Glu Ser Val Ser Tyr Leu Phe Asn Asn Ala Val Met
219 305          310          315
223 <210> SEQ ID NO: 3
224 <211> LENGTH: 5369
225 <212> TYPE: DNA
226 <213> ORGANISM: Ashbya gossypii
228 <220> FEATURE:
229 <221> NAME/KEY: CDS
230 <222> LOCATION: 55..1482
232 <220> FEATURE:
233 <221> NAME/KEY: CDS
234 <222> LOCATION: 1767..3299
236 <220> FEATURE:
237 <221> NAME/KEY: CDS
238 <222> LOCATION: 3588..4703
240 <400> SEQUENCE: 3
242 aagcttgacc ttggctggca cttgagtcgg cagacaggtg gactaaccgc agca atg      57
243                                     Met
244                                     1
246 gat cgt ggt tgt aaa ggt atc tct tat gtg ctc agt gca atg gtt ttt      105
247 Asp Arg Gly Cys Lys Gly Ile Ser Tyr Val Leu Ser Ala Met Val Phe
248          5          10          15
250 cac ata ata ccg att aca ttt gaa ata tcg atg gta tgt ggc ata ttg      153
251 His Ile Ile Pro Ile Thr Phe Glu Ile Ser Met Val Cys Gly Ile Leu
252          20          25          30
254 aca tac cag ttt ggt gct tcc ttc gct gct ata aca ttc tcg act atg      201
255 Thr Tyr Gln Phe Gly Ala Ser Phe Ala Ala Ile Thr Phe Ser Thr Met
256          35          40          45
258 ctt ctt tac tcc atc ttt act ttc aga acg acg gcg tgg cgc aca cgg      249
259 Leu Leu Tyr Ser Ile Phe Thr Phe Arg Thr Thr Ala Trp Arg Thr Arg
260          50          55          60          65
262 ttt agg cgt gat gcg aac aag gct gac aat aag gcc gct agt gtg gca      297
263 Phe Arg Arg Asp Ala Asn Lys Ala Asp Asn Lys Ala Ala Ser Val Ala
264          70          75          80
266 ttg gat tcc cta ata aat ttt gaa gct gta aag tat ttc aat aac gag      345
267 Leu Asp Ser Leu Ile Asn Phe Glu Ala Val Lys Tyr Phe Asn Asn Glu
268          85          90          95
270 aag tac ctt gcg gac aag tat cac aca tcc ttg atg aag tac cgg gat      393
271 Lys Tyr Leu Ala Asp Lys Tyr His Thr Ser Leu Met Lys Tyr Arg Asp
272          100          105          110
274 tcc cag ata aag gtc tcg caa tcg ctg gcg ttt ttg aac acc ggc cag      441
275 Ser Gln Ile Lys Val Ser Gln Ser Leu Ala Phe Leu Asn Thr Gly Gln
276          115          120          125
278 aac cta att ttt acc act gca ctg act gca atg atg tat atg gcc tgt      489
279 Asn Leu Ile Phe Thr Thr Ala Leu Thr Ala Met Met Tyr Met Ala Cys

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/076,157

DATE: 04/16/2002

TIME: 14:33:26

Input Set : A:\EP.txt

Output Set: N:\CRF3\04162002\J076157.raw

280	130		135		140		145		
282	aat	ggt	gtt	atg	cag	ggc	tct	ctt	aca
283	Asn	Gly	Val	Met	Gln	Gly	Ser	Leu	Thr
284									
286	aat	caa	ctg	gta	ttc	cag	ctc	tcc	gtg
287	Asn	Gln	Leu	Val	Phe	Gln	Leu	Ser	Val
288									
290	gtc	tac	cgt	gat	ctc	aag	cag	tct	ctg
291	Val	Tyr	Arg	Asp	Leu	Lys	Gln	Ser	Leu
292									
294	aaa	ctg	caa	aaa	aat	cag	gtc	aca	att
295	Lys	Leu	Gln	Lys	Asn	Gln	Val	Thr	Ile
296									
298	aac	cta	cca	ata	cac	aaa	ccg	ttg	gat
299	Asn	Leu	Pro	Ile	His	Lys	Pro	Leu	Asp
300	210								
302	ttt	ggc	tat	gac	ccg	gag	cgg	cgt	ata
303	Phe	Gly	Tyr	Asp	Pro	Glu	Arg	Arg	Ile
304									
306	atc	cca	gct	gga	atg	aag	act	gcc	ata
307	Ile	Pro	Ala	Gly	Met	Lys	Thr	Ala	Ile
308									
310	aag	tcc	acc	att	ttg	aag	ctc	gta	ttt
311	Lys	Ser	Thr	Ile	Leu	Lys	Leu	Val	Phe
312									
314	ggt	cgt	atc	cta	gtt	ggc	ggc	aca	gat
315	Gly	Arg	Ile	Leu	Val	Gly	Gly	Thr	Asp
316									
318	tct	tta	cgg	aag	gct	atc	ggt	gtc	gtg
319	Ser	Leu	Arg	Lys	Ala	Ile	Gly	Val	Val
320	290								
322	aat	gac	aca	atc	tgg	gag	aag	gtt	aaa
323	Asn	Asp	Thr	Ile	Trp	Glu	Asn	Val	Lys
324									
326	gac	gat	gag	att	ctc	agg	gcc	ata	gaa
327	Asp	Asp	Glu	Ile	Leu	Arg	Ala	Ile	Glu
328									
330	ctc	cag	aac	cta	cca	aag	ggc	gct	tcc
331	Leu	Gln	Asn	Leu	Pro	Lys	Gly	Ala	Ser
332									
334	ttg	atg	atc	agc	gga	ggt	gag	aaa	caa
335	Leu	Met	Ile	Ser	Gly	Gly	Glu	Lys	Gln
336									
338	ctt	ttg	aag	gac	gct	ccg	ctg	atg	ttt
339	Leu	Leu	Lys	Asp	Ala	Pro	Leu	Met	Phe
340	370								
342	ctg	gat	aca	cac	aca	gag	cag	gca	ctc
343	Leu	Asp	Thr	His	Thr	Glu	Gln	Ala	Leu
344									

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/076,157

DATE: 04/16/2002

TIME: 14:33:28

Input Set : A:\EP.txt

Output Set: N:\CRF3\04162002\J076157.raw